

# DISCRETIZATION AND AFFINE APPROXIMATION IN HIGH DIMENSIONS

SEAN LI AND ASSAF NAOR

**ABSTRACT.** Lower estimates are obtained for the macroscopic scale of affine approximability of vector-valued Lipschitz functions on finite dimensional normed spaces, completing the work of Bates, Johnson, Lindenstrass, Preiss and Schechtman. This yields a new approach to Bourgain's discretization theorem for superreflexive targets.

## 1. INTRODUCTION

Let  $X, Y$  be Banach space with (closed) unit balls  $B_X, B_Y$ , respectively. For  $\varepsilon \in (0, 1)$  define  $r^{X \rightarrow Y}(\varepsilon)$  to be the supremum over those  $r \in (0, 1]$  such that for every Lipschitz function  $f : B_X \rightarrow Y$  there exists  $y \in X$  and  $\rho \in [r, \infty)$  such that  $y + \rho B_X \subseteq B_X$ , and there exists an affine mapping  $A : X \rightarrow Y$  satisfying

$$\sup_{z \in y + \rho B_X} \frac{\|f(z) - A(z)\|}{\rho} \leq \varepsilon \|f\|_{\text{Lip}}, \quad (1)$$

where  $\|f\|_{\text{Lip}}$  is the Lipschitz constant of  $f$ . If no such  $r \in (0, 1]$  exists then set  $r^{X \rightarrow Y}(\varepsilon) = 0$ . We call  $r^{X \rightarrow Y}(\cdot)$  the **modulus of affine approximability** corresponding to  $X, Y$ .

The assertion  $r^{X \rightarrow Y}(\varepsilon) \geq r$  means that every Lipschitz function on the unit ball of  $X$  is  $\varepsilon$ -close (after appropriate normalization) to an affine function on some sub-ball of radius  $r$ . Thus, while a differentiation statement corresponds to an assertion about the infinitesimal regularity of a function, bounding  $r^{X \rightarrow Y}(\varepsilon)$  from below corresponds to proving a quantitative differentiation theorem about the regularity of Lipschitz functions on a macroscopic scale. This statement isn't quite precise, since there is no requirement of the affine mapping  $A$  in (1) to have any relation to the derivative of  $f$  at  $y$ , but it turns out that for interesting applications it suffices (and necessary) to allow for an arbitrary affine approximation of  $f$ .

Bates, Johnson, Lindenstrass, Preiss and Schechtman introduced the above affine approximability problem in [2], where it was shown to have applications to the theory of nonlinear quotient mappings. Following [2] we say that the space of Lipschitz mappings

$$\text{Lip}(X, Y) \stackrel{\text{def}}{=} \{f : X \rightarrow Y : \|f\|_{\text{Lip}} < \infty\}$$

has the Uniform Approximation by Affine Property (UAAP) if  $r^{X \rightarrow Y}(\varepsilon) > 0$  for all  $\varepsilon \in (0, 1)$ . A beautiful theorem of [2] asserts that  $\text{Lip}(X, Y)$  has the UAAP if and only if one of the spaces  $\{X, Y\}$  is finite dimensional, and the other space is superreflexive. Recall that a Banach space  $Z$  is superreflexive if any Banach space that is finitely representable in  $Z$

---

S. L. was supported by NSF grant CCF-0832795. A. N. was supported by NSF grant CCF-0832795, BSF grant 2006009, and the Packard Foundation. Some of this work was completed when both authors were in residence at the MSRI Quantitative Geometry program.

must be reflexive; equivalently all the ultrapowers of  $Z$  are reflexive<sup>1</sup>. Due to deep works of James [16, 17], Enflo [9] and Pisier [21], we know that  $Z$  is superreflexive if and only if it admits an equivalent norm  $\|\cdot\|$  for which there exist  $p \in [2, \infty)$  and  $K \in [1, \infty)$  such that

$$\forall x, y \in Z, \quad 2\|x\|^p + \frac{2}{K^p}\|y\|^p \leq \|x + y\|^p + \|x - y\|^p. \quad (2)$$

A norm that satisfies (2) is said to be uniformly convex of power type  $p$ ; readers who are not familiar with the theory of superreflexivity can take the above renorming statement as the definition of superreflexivity. For concreteness, we recall [11, 1] that for  $q \in (1, \infty)$  the usual norm on an  $L_q(\mu)$  space satisfies (2) with  $p = \max\{q, 2\}$  and  $K = \max\{1/\sqrt{q-1}, 1\}$ .

Assume from now on that  $n = \dim X < \infty$  and  $Y$  is superreflexive. The theorem of Bates, Johnson, Lindenstrass, Preiss and Schechtman says that  $r^{X \rightarrow Y}(\varepsilon), r^{Y \rightarrow X}(\varepsilon) > 0$  for every  $\varepsilon \in (0, 1)$ . The proof in [2] of  $r^{Y \rightarrow X}(\varepsilon) > 0$  is effective, yielding a concrete lower bound on  $r^{Y \rightarrow X}(\varepsilon)$ . This lower bound is quite small: a  $O(n)$ -fold iterated exponential of  $-1/\varepsilon$  and geometric parameters that measure the degree to which  $Y$  is superreflexive. We leave the investigation of the true asymptotic behavior of  $r^{Y \rightarrow X}(\varepsilon)$  as an interesting open problem.

Our main purpose here is to obtain a concrete lower bound on  $r^{X \rightarrow Y}(\varepsilon)$ . While we are partly motivated by an application of such bounds to Bourgain's discretization problem, as will be described in Section 1.1, our main motivation is that the proof in [2] of the estimate  $r^{X \rightarrow Y}(\varepsilon) > 0$  proceeds by contradiction using an ultrapower argument, and as such it does not yield any concrete quantitative information on  $r^{X \rightarrow Y}(\varepsilon)$ .

We briefly recall the argument of [2]. The contrapositive assumption  $r^{X \rightarrow Y}(\varepsilon) = 0$  means that for every  $k \in \mathbb{N}$  there is a 1-Lipschitz function  $f_k : B_X \rightarrow Y$  with  $f_k(0) = 0$  such that for all balls  $y + rB_X \subseteq B_X$  with  $r \geq 1/k$  and for all affine mappings  $A : X \rightarrow Y$  we have  $\|f_k - A\|_{L_\infty(y+rB_X)} \geq \varepsilon r$ . Let  $\mathcal{U}$  be a free ultrafilter on  $\mathbb{N}$  and consider the mapping  $f_{\mathcal{U}} : B_X \rightarrow Y_{\mathcal{U}}$  given by  $f_{\mathcal{U}}(x) = (f_k(x))_{k=1}^\infty$ . Here  $Y_{\mathcal{U}}$  denotes the ultrapower of  $Y$ ; since  $Y$  is superreflexive we are ensured that  $Y_{\mathcal{U}}$  is reflexive. A moment of thought reveals that  $f_{\mathcal{U}}$  is 1-Lipschitz yet it cannot have a point of differentiability. This contradicts the fact [12] that reflexive spaces have the Radon-Nikodým property (see [4, Ch. 5]), and hence  $f_{\mathcal{U}}$  is differentiable almost everywhere. Due to the ineffectiveness of this argument, the estimation of the fundamental parameter  $r^{X \rightarrow Y}(\varepsilon)$  is a basic question that [2] left open. This problem is resolved here via the following theorem.

**Theorem 1.1.** *Fix  $n \in \mathbb{N}$ ,  $p \in [2, \infty)$  and  $K \in [1, \infty)$ . Assume that  $n = \dim X < \infty$  and the norm of  $Y$  satisfies (2). Then for all  $\varepsilon \in (0, \frac{1}{2})$  we have*

$$r^{X \rightarrow Y}(\varepsilon) \geq \begin{cases} \varepsilon^{(16K/\varepsilon)^p} & \text{if } n = 1, \\ \varepsilon^{K^p n^{20(n+p)}/\varepsilon^{2p+2n-2}} & \text{if } n \geq 2. \end{cases} \quad (3)$$

In Section 4 we present an example showing that if  $\varepsilon \in (0, \frac{1}{2})$  and  $p \in [2, \infty)$  then for  $X_0 = \ell_2^n$  and  $Y_0 = \ell_2(\ell_p)$  we have

$$r^{X_0 \rightarrow Y_0}(\varepsilon) \leq \frac{1}{\sqrt{n}} e^{-(\kappa/\varepsilon)^p},$$

where  $\kappa \in (0, \infty)$  is a universal constant. Note that  $\ell_2(\ell_p)$  satisfies (2); see [11]. Thus, when  $n = 1$  Theorem 1.1 is quite sharp as  $\varepsilon \rightarrow 0$  (up to a  $\log(1/\varepsilon)$  term in the exponent), but it

---

<sup>1</sup>See [8] for background on finite representability and ultrapowers of Banach spaces.

remains a very interesting open problem to determine the asymptotic behavior of  $r^{X \rightarrow Y}(\varepsilon)$  as  $n \rightarrow \infty$ . It is worthwhile to single out the purely Hilbertian special case of this problem.

**Question 1.** *What is the asymptotic behavior of  $r^{\ell_2^n \rightarrow \ell_2}(\frac{1}{2})$  as  $n \rightarrow \infty$ ?*

We note that despite the fact that the gap between (3) and the above upper bound on  $r^{X \rightarrow Y}(\varepsilon)$  is very large as  $n \rightarrow \infty$ , the lower estimate on  $r^{X \rightarrow Y}(\varepsilon)$  in (3) is sufficiently strong to match the best-known bound in Bourgain's discretization theorem for superreflexive targets; see Section 1.1.

In our forthcoming article [15], written jointly with Tuomas Hytönen, we study a natural variant of the UAAP by replacing the  $L_\infty$  requirement in (1) by

$$\left( \frac{1}{\rho^n \text{vol}(B_X)} \int_{y+\rho B_X} \left( \frac{\|f(z) - A(z)\|}{\rho} \right)^p dz \right)^{1/p} \leq \varepsilon \|f\|_{\text{Lip}}.$$

In this setting, we obtain in [15] asymptotically stronger lower bounds on  $\rho$  when  $Y$  is a UMD Banach space (see [6] for a detailed discussion of UMD spaces). Unlike our proof of Theorem 1.1, which is entirely geometric, the arguments in [15] are based on vector-valued Littlewood-Paley theory.

**1.1. Bourgain's discretization theorem.** Let  $(X, d_X)$  and  $(Y, d_Y)$  be metric spaces. The distortion of  $X$  in  $Y$ , denoted  $c_Y(X)$ , is the infimum over those  $D \in [1, \infty]$  for which there exists  $f : X \rightarrow Y$  and  $s \in (0, \infty)$  satisfying

$$\forall x, y \in X, \quad sd_X(x, y) \leq d_Y(f(x), f(y)) \leq Dsd_X(x, y).$$

Suppose now that  $(X, \|\cdot\|_X)$  and  $(Y, \|\cdot\|_Y)$  are normed spaces with  $\dim(X) < \infty$  and  $\dim(Y) = \infty$ . For  $\varepsilon \in [0, 1]$  let  $\delta_{X \hookrightarrow Y}(\varepsilon)$  be the supremum over those  $\delta \in (0, 1)$  such that every  $\delta$ -net  $\mathcal{N}_\delta$  of  $B_X$  satisfied  $c_Y(\mathcal{N}_\delta) \geq (1 - \varepsilon)c_Y(X)$ .

A classical theorem of Ribe asserts that  $\delta_{X \hookrightarrow Y}(\varepsilon) > 0$  for all  $\varepsilon \in (0, 1)$ . A different proof of this fact, due to Heinrich and Mankiewicz, was obtained in [14]. Bourgain [5] discovered yet another proof of the positivity of  $\delta_{X \hookrightarrow Y}(\varepsilon)$ , which, unlike previous proofs, yields the following concrete estimate, known as *Bourgain's discretization theorem*.

$$\delta_{X \hookrightarrow Y}(\varepsilon) \geq e^{-(n/\varepsilon)^{O(n)}}. \quad (4)$$

It is an intriguing open question to determine the asymptotic behavior of the best possible lower bound on  $\inf\{\delta_{X \hookrightarrow Y}(\varepsilon) : \dim(X) = n\}$ . This is of interest even for special classes of normed spaces  $Y$ , though there has been little progress on this problem besides the improved estimate  $\delta_{X \hookrightarrow L_p}(\varepsilon) \gtrsim \varepsilon^2/n^{5/2}$ , which was obtained in [13] (here  $p \in [1, \infty)$  and the implied constant is independent of  $p$ ).

We shall now describe a different approach to Bourgain's discretization theorem based on Theorem 1.1. If an affine mapping is bi-Lipschitz on a fine enough net of a ball  $y + \rho B_X$  then it is also bi-Lipschitz on all of  $X$ . It is therefore natural to approach the problem of estimating  $\delta_{X \hookrightarrow Y}(\varepsilon)$  by first extending the embedding of the net  $\mathcal{N}_\delta$  to a Lipschitz function defined on all of  $X$ , and then finding a large enough ball on which the extended function is approximately affine. By the theorem of Bates, Johnson, Lindensrauss, Preiss and Schechtman, for this strategy to work we need  $Y$  to be superreflexive. Bourgain's discretization theorem is interesting even for superreflexive targets, and moreover the estimate (4) is the best known estimate even with this additional restriction on  $Y$ . It turns out that the above strategy,

when combined with our estimate (3), suffices to match Bourgain's bound (4) when  $Y$  is superreflexive. The details of this link between Theorem 1.1 and (4) are explained below.

Fix  $\varepsilon \in (0, 1)$ ,  $p \in [2, \infty)$  and  $K \in [1, \infty)$ . Suppose that  $\dim(X) = n \geq 2$  and the norm of  $Y$  satisfies the uniform convexity condition (2). Set

$$\delta = e^{-K^p(n/\varepsilon)^{C(n+p)}}, \quad (5)$$

where  $C \in (1, \infty)$  is a universal constant that will be determined later.

Let  $\mathcal{N}_\delta$  be a  $\delta$ -net of  $B_X$  and write  $D = c_Y(\mathcal{N}_\delta)$ . Note that, by John's theorem [18], we have the a priori bound  $D \leq n$ . Take  $f : \mathcal{N}_\delta \rightarrow Y$  satisfying

$$\forall x, y \in \mathcal{N}_\delta, \quad \|x - y\|_X \leq \|f(x) - f(y)\|_Y \leq \left(1 + \frac{\varepsilon}{16}\right) D \|x - y\|_X. \quad (6)$$

By a Lipschitz extension theorem of Johnson, Lindenstrauss and Schechtman [19], there exists  $F : X \rightarrow Y$  that coincides with  $f$  when restricted to  $\mathcal{N}_\delta$ , and  $\|F\|_{\text{Lip}} \leq cnD$ , where  $c \in (1, \infty)$  is a universal constant.

By Theorem 1.1 there exist  $y \in X$ ,  $z \in Y$ , a linear mapping  $T : X \rightarrow Y$ , and a radius

$$\rho \geq \varepsilon^{K^p n^{20(n+p)} (32cnD/\varepsilon)^{2p+2n-2}}. \quad (7)$$

such that  $y + \rho B_X \subseteq B_X$  and

$$\forall x \in y + \rho B_X, \quad \|F(x) - z - Tx\|_Y \leq \frac{\varepsilon}{32} \rho. \quad (8)$$

Note that it follows from (7) that we can choose the constant  $C$  in (5) so that

$$\rho \geq \frac{64n\delta}{\varepsilon}. \quad (9)$$

Fix  $u \in X$  with  $\|u\|_X = 1$ . Choose  $v, w \in \mathcal{N}_\delta \cap (y + \rho B_X)$  such that  $\|v - y\|_X \leq \delta$  and  $\|w - y - \frac{\rho}{2}u\|_X \leq \delta$ . Thus  $\|w - v - \frac{\rho}{2}u\|_X \leq 2\delta$ , and consequently  $\|w - v\|_X \in [\rho/2 - 2\delta, \rho/2 + 2\delta]$ . Using the fact that  $F$  extends  $f$ ,

$$\begin{aligned} \|Tw - Tv\|_Y &\stackrel{(8)}{\leq} \|f(w) - f(v)\|_Y + \frac{\varepsilon\rho}{16} \stackrel{(6)}{\leq} \left(1 + \frac{\varepsilon}{16}\right) D \|w - v\|_X + \frac{\varepsilon\rho}{16} \\ &\leq \left(1 + \frac{\varepsilon}{16}\right) D \left(\frac{\rho}{2} + 2\delta\right) + \frac{\varepsilon\rho}{16} \stackrel{(9)}{\leq} \left(1 + \frac{\varepsilon}{4}\right) \frac{\rho}{2} D. \end{aligned}$$

Hence  $\|Tu\|_Y \leq \frac{2}{\rho} \|Tw - Tv\|_Y + \frac{2\|T\|}{\rho} \|w - v - \frac{\rho}{2}u\|_X \leq \left(1 + \frac{\varepsilon}{4}\right) D + \frac{4\delta\|T\|}{\rho}$ . Since this holds for all unit vectors  $u \in X$ ,

$$\|T\| \leq \frac{1 + \varepsilon/4}{1 - 4\delta/\rho} D \leq \left(1 + \frac{\varepsilon}{2}\right) D \leq 2n. \quad (10)$$

Now,

$$\|Tw - Tv\|_Y \stackrel{(8)}{\geq} \|f(w) - f(v)\|_Y - \frac{\varepsilon\rho}{16} \stackrel{(6)}{\geq} \|w - v\|_X - \frac{\varepsilon\rho}{16} \geq \frac{\rho}{2} - 2\delta - \frac{\varepsilon\rho}{16} \geq \left(1 - \frac{\varepsilon}{4}\right) \frac{\rho}{2}.$$

Hence,

$$\|Tu\|_Y \geq \frac{2}{\rho} \|Tw - Tv\|_Y - \frac{2\|T\|}{\rho} \left\|w - v - \frac{\rho}{2}u\right\|_X \stackrel{(10)}{\geq} 1 - \frac{\varepsilon}{4} - \frac{8n\delta}{\rho} \stackrel{(9)}{\geq} 1 - \frac{\varepsilon}{2}.$$

We have proved that  $c_Y(X) \leq \frac{1+\varepsilon/2}{1-\varepsilon/2}D = \frac{1+\varepsilon/2}{1-\varepsilon/2}c_Y(\mathcal{N}_\delta) \leq \frac{1}{1-\varepsilon}c_Y(\mathcal{N}_\delta)$ . Thus, recalling the choice of  $\delta$  in (5),

$$\delta_{X \hookrightarrow Y}(\varepsilon) \geq e^{-K^p(n/\varepsilon)^{C(n+p)}}. \quad (11)$$

**Remark 1.1.** In fact, we have the general estimate

$$\delta_{X \hookrightarrow Y}(\varepsilon) \geq \frac{\varepsilon}{n} \cdot r^{X \rightarrow Y} \left( \frac{\kappa \varepsilon}{c_Y(X)} \right), \quad (12)$$

where  $\kappa \in (0, \infty)$  is a universal constant. This estimate follows from a more careful application of the above reasoning. Specifically, we used the Lipschitz extension theorem of Johnson, Lindenstrauss and Schechtman [19] to obtain the function  $F$ . This theorem ignores the fact that  $f$  was defined on a  $\delta$ -net: it would apply equally well if  $f$  were defined on *any* subset of  $X$ . One can exploit the additional information that the domain of  $f$  is a net by invoking an approximate Lipschitz extension theorem of Bourgain [5]. This theorem states that for every  $\tau \in (20\delta, 1)$  there exists a function  $F_\tau : X \rightarrow Y$  such that  $\|F_\tau(x) - f(x)\|_Y \leq \tau$  for every  $x \in \mathcal{N}_\delta$  and the Lipschitz constant of  $F_\tau$  on  $\frac{1}{2}B_X$  is  $(1 + O(n\delta/\tau))(1 + \varepsilon/16)D$  (this formulation of Bourgain’s approximate extension theorem is not stated explicitly in [5], but it easily follows from the argument in [13, Sec. 3]). Now, one can deduce (12) by repeating *mutatis mutandis* the above proof while optimizing over  $\tau$ . We omit the details since the resulting estimate, when applied to our bounds (3), only affects the constant  $C$  in (11).

## 2. PROOF OF THEOREM 1.1 WHEN $n = 1$

The proof of Theorem 1.1 when  $n = 1$  follows well-established metric differentiation methodology. This type of reasoning, also known as the approximate midpoint argument, seems to have been first used by Enflo in his classical proof that  $L_1$  and  $\ell_1$  are not uniformly homeomorphic; see [3]. The basic idea is that a Lipschitz function  $f : \mathbb{R} \rightarrow Y$  must map the midpoints between many pairs of points  $x, y \in \mathbb{R}$  to “almost midpoints” of  $f(x)$  and  $f(y)$ . See Chapter 10 of [4] for a precise formulation of this principle. One can iterate this idea to deduce that  $f$  must map many “discretized geodesic segments” to “discretized almost geodesics”. Such an iteration of the midpoint argument is contained in e.g. [20, Prop. 1.4.9], and a striking recent application of this type of reasoning can be found in [10]. When the target space is uniformly convex, approximate geodesics must be close to straight lines. This rigidity statement explains why one can hope to find a macroscopically large region on which  $f$  is almost affine. In order to obtain good quantitative control on the size of such a region one follows the general strategy that is explained in Appendix 2 of [7]. Using the terminology of [7], the “coercive quantity” in our setting is the functional  $E_m^{a,b}(\cdot)$  defined below.

In Section 3 we build on the tools developed in this section to deduce Theorem 1.1 when  $n \geq 2$ . In this setting one must deal with higher-dimensional phenomena, and to obtain good dimension-dependent bounds. We will argue that there must exist a cube on which a given Lipschitz function maps all axis-parallel discretized line segments to almost-straight lines. This does not imply that the function itself is almost affine on the same cube: at best it means that it is almost multi-linear. Therefore an additional argument is needed in order to find a scale on which the function is almost affine. Moreover, to obtain good control on this scale we reason about a more complicated (two-parameter) coercive quantity; see (42).

**Lemma 2.1.** Fix  $p \in [2, \infty)$ . Suppose that  $(Y, \|\cdot\|_Y)$  is a Banach space satisfying the uniform convexity condition (2). Fix  $a, b \in \mathbb{R}$  with  $a < b$  and  $h : [a, b] \rightarrow Y$ . For  $m \in \mathbb{N} \cup \{0\}$  define

$$E_m^{a,b}(h) \stackrel{\text{def}}{=} \frac{1}{2^m} \sum_{k=0}^{2^m-1} \left\| \frac{h(a + (k+1)2^{-m}(b-a)) - h(a + k2^{-m}(b-a))}{2^{-m}(b-a)} \right\|_Y^p. \quad (13)$$

Then

$$E_m^{a,b}(h) \geq \frac{\|h(b) - h(a)\|_Y^p}{(b-a)^p} + \frac{1}{(2K)^p} \max_{k \in \{0, \dots, 2^m\}} \frac{\left\| h\left(a + \frac{k}{2^m}(b-a)\right) - L_h^{a,b}\left(a + \frac{k}{2^m}(b-a)\right) \right\|_Y^p}{(b-a)^p},$$

where  $K \in (0, \infty)$  is the constant appearing in (2) and  $L_h^{a,b}(h) : [a, b] \rightarrow Y$  is the linear interpolation of the values of  $h$  on the endpoints of the interval  $[a, b]$ , i.e.,

$$\forall t \in \mathbb{R}, \quad L_h^{a,b}(t) \stackrel{\text{def}}{=} \frac{t-a}{b-a}h(b) + \frac{b-t}{b-a}h(a). \quad (14)$$

*Proof.* We may assume without loss of generality that  $a = 0$  and  $b = 1$ . In this case, denote for the sake of simplicity  $E_m^{0,1}(h) = E_m(h)$  and  $L_h^{0,1} = L_h$ . We will actually prove the following slightly stronger statement by induction on  $m$ : for every  $k \in \{0, \dots, 2^m\}$ ,

$$E_m(h) \geq \|h(1) - h(0)\|_Y^p + \frac{2^p}{K^p (3 - 2^{-(m-1)})^{p-1}} \cdot \left\| h\left(\frac{k}{2^m}\right) - L_h\left(\frac{k}{2^m}\right) \right\|_Y^p. \quad (15)$$

Since  $E_0(h) = \|h(1) - h(0)\|_Y^p$ , the desired inequality (15) holds as equality when  $m = 0$ . Fix  $m \in \mathbb{N}$  and assume that (15) holds true with  $m$  replaced by  $m-1$ .

Convexity of  $\|\cdot\|_Y^p$  implies that for every  $m \in \mathbb{N}$ ,

$$\begin{aligned} E_m(h) &= 2^{m(p-1)} \sum_{k=0}^{2^m-1} \left\| h\left(\frac{k+1}{2^m}\right) - h\left(\frac{k}{2^m}\right) \right\|_Y^p \\ &= 2^{m(p-1)+1} \sum_{j=0}^{2^{m-1}-1} \frac{\left\| h\left(\frac{2j+1}{2^m}\right) - h\left(\frac{2j}{2^m}\right) \right\|_Y^p + \left\| h\left(\frac{2j+2}{2^m}\right) - h\left(\frac{2j+1}{2^m}\right) \right\|_Y^p}{2} \\ &\geq 2^{m(p-1)+1} \sum_{j=0}^{2^{m-1}-1} \left\| \frac{h\left(\frac{j+1}{2^{m-1}}\right) - h\left(\frac{j}{2^{m-1}}\right)}{2} \right\|_Y^p \\ &= E_{m-1}(h). \end{aligned} \quad (16)$$

Hence, if  $k \in \{0, \dots, 2^m\}$  is even then by the inductive hypothesis

$$\begin{aligned} E_m(h) &\stackrel{(16)}{\geq} E_{m-1}(h) \stackrel{(15)}{\geq} E_0(h) + \frac{2^p}{K^p (3 - 2^{-(m-2)})^{p-1}} \left\| h\left(\frac{k/2}{2^{m-1}}\right) - L_h\left(\frac{k/2}{2^{m-1}}\right) \right\|_Y^p \\ &\geq E_0(h) + \frac{2^p}{K^p (3 - 2^{-(m-1)})^{p-1}} \left\| h\left(\frac{k}{2^m}\right) - L_h\left(\frac{k}{2^m}\right) \right\|_Y^p. \end{aligned}$$

It therefore suffices to prove (15) when  $k$  is odd, say,  $k = 2j + 1$ . In this case, by reasoning analogously to (16), we see that

$$\begin{aligned}
& \frac{E_m(h) - E_{m-1}(h)}{2^{m(p-1)+1}} \\
& \geq \frac{\left\| h\left(\frac{2j+1}{2^m}\right) - h\left(\frac{2j}{2^m}\right) \right\|_Y^p + \left\| h\left(\frac{2j+2}{2^m}\right) - h\left(\frac{2j+1}{2^m}\right) \right\|_Y^p}{2} - \left\| \frac{h\left(\frac{2j+2}{2^m}\right) - h\left(\frac{2j}{2^m}\right)}{2} \right\|_Y^p \\
& \geq \frac{1}{K^p} \left\| \frac{h\left(\frac{j}{2^{m-1}}\right) + h\left(\frac{j+1}{2^{m-1}}\right)}{2} - h\left(\frac{k}{2^m}\right) \right\|_Y^p,
\end{aligned} \tag{17}$$

where in (17) we used (2) with

$$x = \frac{h\left(\frac{2j+2}{2^m}\right) - h\left(\frac{2j}{2^m}\right)}{2} \quad \text{and} \quad y = h\left(\frac{2j+1}{2^m}\right) - \frac{h\left(\frac{2j}{2^m}\right) + h\left(\frac{2j+2}{2^m}\right)}{2}.$$

The inductive hypothesis implies that

$$\begin{aligned}
& \frac{K^p}{2^p} (E_{m-1}(h) - E_0(h)) \\
& \geq \frac{\max \left\{ \left\| h\left(\frac{j}{2^{m-1}}\right) - L_h\left(\frac{j}{2^{m-1}}\right) \right\|_Y^p, \left\| h\left(\frac{j+1}{2^{m-1}}\right) - L_h\left(\frac{j+1}{2^{m-1}}\right) \right\|_Y^p \right\}}{(3 - 2^{-(m-2)})^{p-1}}.
\end{aligned} \tag{18}$$

Since  $L_h$  is affine, by convexity of  $\|\cdot\|_Y^p$  we have

$$\begin{aligned}
& \left\| \frac{h\left(\frac{j}{2^{m-1}}\right) + h\left(\frac{j+1}{2^{m-1}}\right)}{2} - L_h\left(\frac{k}{2^m}\right) \right\|_Y^p = \left\| \frac{h\left(\frac{j}{2^{m-1}}\right) - L_h\left(\frac{j}{2^{m-1}}\right) + h\left(\frac{j+1}{2^{m-1}}\right) - L_h\left(\frac{j+1}{2^{m-1}}\right)}{2} \right\|_Y^p \\
& \leq \frac{\left\| h\left(\frac{j}{2^{m-1}}\right) - L_h\left(\frac{j}{2^{m-1}}\right) \right\|_Y^p + \left\| h\left(\frac{j+1}{2^{m-1}}\right) - L_h\left(\frac{j+1}{2^{m-1}}\right) \right\|_Y^p}{2} \\
& \leq \max \left\{ \left\| h\left(\frac{j}{2^{m-1}}\right) - L_h\left(\frac{j}{2^{m-1}}\right) \right\|_Y^p, \left\| h\left(\frac{j+1}{2^{m-1}}\right) - L_h\left(\frac{j+1}{2^{m-1}}\right) \right\|_Y^p \right\}.
\end{aligned} \tag{19}$$

Therefore, using (17), (18) and (19), we have

$$\begin{aligned}
& \frac{K^p}{2^p} (E_m(h) - E_0(h)) \\
& \geq \frac{\left\| \frac{h\left(\frac{j}{2^{m-1}}\right) + h\left(\frac{j+1}{2^{m-1}}\right)}{2} - L_h\left(\frac{k}{2^m}\right) \right\|_Y^p}{(3 - 2^{-(m-2)})^{p-1}} + 2^{(m-1)(p-1)} \left\| \frac{h\left(\frac{j}{2^{m-1}}\right) + h\left(\frac{j+1}{2^{m-1}}\right)}{2} - h\left(\frac{k}{2^m}\right) \right\|_Y^p \\
& \geq \frac{\left( \left\| \frac{h\left(\frac{j}{2^{m-1}}\right) + h\left(\frac{j+1}{2^{m-1}}\right)}{2} - L_h\left(\frac{k}{2^m}\right) \right\|_Y + \left\| \frac{h\left(\frac{j}{2^{m-1}}\right) + h\left(\frac{j+1}{2^{m-1}}\right)}{2} - h\left(\frac{k}{2^m}\right) \right\|_Y \right)^p}{(3 - 2^{-(m-2)} + 2^{-(m-1)})^{p-1}}
\end{aligned} \tag{20}$$

$$\geq \frac{\left\| h\left(\frac{k}{2^m}\right) - L_h\left(\frac{k}{2^m}\right) \right\|_Y^p}{(3 - 2^{-(m-1)})^{p-1}}, \tag{21}$$

where in (20) we used the inequality

$$\forall \alpha, \beta, u, v \in (0, \infty), \quad \frac{u^p}{\alpha^{p-1}} + \frac{v^p}{\beta^{p-1}} \geq \frac{(u+v)^p}{(\alpha+\beta)^{p-1}},$$

which is an immediate consequence of Hölder's inequality. Since inequality (21) is the same as the desired inequality (15), the proof of Lemma 2.1 is complete.  $\square$

*Proof of Theorem 1.1 when  $n = 1$ .* Our goal is to show that if  $(Y, \|\cdot\|_Y)$  is a Banach space satisfying (2) then for every  $\varepsilon \in (0, \frac{1}{2})$  we have

$$r^{\mathbb{R} \rightarrow Y}(\varepsilon) \geq \left(\frac{\varepsilon}{8}\right)^{(8K/\varepsilon)^p}. \quad (22)$$

The fact that (22) is better than the desired estimate (3) is a simple elementary inequality (recall that  $p \geq 2$ ,  $K \geq 1$  and  $0 < \varepsilon < 1/2$ ).

Assume for contradiction that (22) fails. Then there exists  $\varepsilon \in (0, \frac{1}{2})$  and a 1-Lipschitz function  $h : [-1, 1] \rightarrow Y$  such that for every  $-1 \leq a < b \leq 1$  with  $b - a \geq (\varepsilon/8)^{(8K/\varepsilon)^p}$  there exists  $t \in [a, b]$  satisfying  $\|h(t) - L_h^{a,b}(t)\|_Y > \varepsilon(b-a)/2$ . Choose  $m \in \mathbb{N}$  such that  $\varepsilon/8 \leq 2^{-m} < \varepsilon/4$  and take  $k \in \{1, \dots, m\}$  such that if we set  $s = a + k2^{-m}(b-a)$  then  $|s-t| \leq (b-a)/2^{m+1}$ . Because  $f$  is 1-Lipschitz, it follows immediately from the definition (14) of  $L_h^{a,b}$  that it is also 1-Lipschitz. Hence,

$$\begin{aligned} \max_{k \in \{0, \dots, 2^m\}} \frac{\left\| h\left(a + \frac{k}{2^m}(b-a)\right) - L_h^{a,b}\left(a + \frac{k}{2^m}(b-a)\right) \right\|_Y^p}{(b-a)^p} &\geq \frac{\|h(s) - L_h^{a,b}(s)\|_Y^p}{(b-a)^p} \\ &\geq \frac{\left( \|h(t) - L_h^{a,b}(t)\|_Y - \|h(t) - h(s)\|_Y - \|L_h^{a,b}(t) - L_h^{a,b}(s)\|_Y \right)^p}{(b-a)^p} \geq \left( \frac{\varepsilon}{2} - \frac{1}{2^m} \right)^p \geq \frac{\varepsilon^p}{4^p}. \end{aligned}$$

Consequently, it follows from Lemma 2.1 that

$$-1 \leq a < b \leq 1 \wedge b - a \geq (\varepsilon/8)^{(8K/\varepsilon)^p} \implies E_m^{a,b}(h) \geq \frac{\|h(b) - h(a)\|_Y^p}{(b-a)^p} + \left( \frac{\varepsilon}{8K} \right)^p. \quad (23)$$

For  $k \in \mathbb{N} \cup \{0\}$  and  $j \in \{0, \dots, 2^{km}\}$  denote  $a_j^k = -1 + j/2^{km-1}$ . If  $1/2^{km-1} \geq (\varepsilon/8)^{(8K/\varepsilon)^p}$  then it follows from (23) that for every  $j \in \{0, \dots, 2^{km} - 1\}$  we have

$$E_m^{a_j^k, a_{j+1}^k}(h) \geq \frac{\|h(a_{j+1}^k) - h(a_j^k)\|_Y^p}{2^{(km-1)p}} + \left( \frac{\varepsilon}{8K} \right)^p. \quad (24)$$

Hence,

$$\begin{aligned} E_{(k+1)m}^{-1,1}(h) &\stackrel{(13)}{=} 2^{-km} \sum_{j=0}^{2^{km}-1} E_m^{a_j^k, a_{j+1}^k}(h) \\ &\stackrel{(24)}{\geq} 2^{-km} \sum_{j=0}^{2^{km}-1} \frac{\|h(a_{j+1}^k) - h(a_j^k)\|_Y^p}{2^{(km-1)p}} + \left( \frac{\varepsilon}{8K} \right)^p \stackrel{(13)}{=} E_{km}^{-1,1}(h) + \left( \frac{\varepsilon}{8K} \right)^p. \end{aligned} \quad (25)$$

Since  $h$  is 1-Lipschitz, the definition (13) implies that  $E_j^{a,b}(h) \leq 1$  for all  $-1 \leq a < b \leq 1$  and  $j \in \mathbb{N}$ . Denote  $M = \lfloor (1 + (8K/\varepsilon)^p \log_2(8/\varepsilon)) / m \rfloor$ . Then (25) holds for every  $k \in \mathbb{N} \cap [0, M]$ . It follows that  $E_{(M+1)m}^{-1,1}(h) \geq E_0^{-1,1}(h) + (M+1)(\varepsilon/(2K))^p \geq (M+1)(\varepsilon/(2K))^p$ . Observe



that the definition of  $M$ , combined with  $2^{-m} \geq \varepsilon/8$ , implies that  $(M+1)(\varepsilon/(2K))^p > 1$ . Thus  $E_{(M+1)m}^{-1,1}(h) > 1$ , a contradiction.  $\square$

### 3. PROOF OF THEOREM 1.1 WHEN $n \geq 2$

Fix  $n \in \mathbb{N}$  and let  $(X, \|\cdot\|_X)$  be an  $n$ -dimensional normed space. Assume also that  $(Y, \|\cdot\|_Y)$  is a Banach space and  $f : X \rightarrow Y$ . By John's theorem [18] there exists a norm  $\|\cdot\|_2$  on  $X$  which is Hilbertian and satisfies  $\|x\|_2 \leq \|x\|_X \leq \sqrt{n}\|x\|_2$  for all  $x \in X$ . Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis with respect to  $\|x\|_2$ . Via the obvious identifications, we may assume below that  $X = \mathbb{R}^n$  and  $\{e_1, \dots, e_n\}$  is the standard coordinate basis.

For  $y \in \mathbb{R}^n$  and  $j \in \{1, \dots, n\}$  define  $f_j^y : \mathbb{R} \rightarrow Y$  by  $f_j^y(t) = f(y + te_j)$ . Also, given  $m \in \mathbb{N}$  and  $j \in \{1, \dots, n\}$  set  $F_j^m = \{z \in \frac{1}{2^m}\{0, \dots, 2^m\}^n : z_j = 0\}$ . For  $x \in \mathbb{R}^n$  and  $\vartheta \in (0, \infty)$  consider the following quantity

$$\begin{aligned} \mathcal{D}_\vartheta^m(f)(x) &\stackrel{\text{def}}{=} \max_{\substack{j \in \{1, \dots, n\} \\ y \in x + \vartheta F_j^m \\ k \in \{0, \dots, 2^m\}}} \frac{\|f(y + \frac{k\vartheta}{2^m}e_j) - f(y) - \frac{k}{2^m}(f(y + \vartheta e_j) - f(y))\|_X}{\vartheta} \\ &\stackrel{(14)}{=} \max_{\substack{j \in \{1, \dots, n\} \\ y \in x + \vartheta F_j^m \\ k \in \{0, \dots, 2^m\}}} \frac{1}{\vartheta} \left\| f_j^y\left(\frac{k}{2^m}\vartheta\right) - L_{f_j^y}^{0, \vartheta}\left(\frac{k}{2^m}\vartheta\right) \right\|_Y. \end{aligned} \quad (26)$$

**Lemma 3.1.** *Fix  $x \in \mathbb{R}^n$ ,  $m \in \mathbb{N}$  and  $\varepsilon, \vartheta \in (0, \infty)$  with  $2^m \geq 2/\varepsilon \geq 10n^2$ . Suppose that  $f : \mathbb{R}^n \rightarrow Y$  satisfies  $\|f(y) - f(z)\| \leq \|y - z\|_2$  for all  $y, z \in x + [0, \vartheta]^n$ , i.e.,  $f$  is 1-Lipschitz with respect to the Euclidean metric on the cube  $x + [0, \vartheta]^n$ . Suppose also that  $\mathcal{D}_\vartheta^m(f)(x) \leq \varepsilon$ . Then there exists an affine mapping  $A : \mathbb{R}^n \rightarrow Y$  such that*

$$\sup_{z \in x + [0, \sqrt{\varepsilon}\vartheta]^n} \|f(z) - A(z)\|_Y \leq 8n^2\varepsilon\vartheta. \quad (27)$$

*Proof.* By translation and rescaling we may assume without loss of generality that  $x = 0$  and  $\vartheta = 1$ . We will prove by induction on  $n$  that there exist vectors  $\{v_S\}_{S \subseteq \{1, \dots, n\}} \subseteq Y$  with

$$v_\emptyset = f(0) \quad \text{and} \quad \forall \emptyset \neq S \subseteq \{1, \dots, n\}, \quad \|v_S\|_Y \leq 2^{|S|-1}, \quad (28)$$

such that for every  $y \in \frac{1}{2^m}\{0, \dots, 2^m\}^n$  we have

$$\left\| f(y) - \sum_{S \subseteq \{1, \dots, n\}} W_S(y) v_S \right\|_Y \leq \varepsilon n, \quad (29)$$

where the Walsh functions  $\{W_S : \mathbb{R}^n \rightarrow \mathbb{R}\}_{S \subseteq \{1, \dots, n\}}$  are defined as usual by  $W_S(y) \stackrel{\text{def}}{=} \prod_{i \in S} y_i$ .

Assuming for the moment that this assertion has been proven, we proceed to deduce (27). Define  $A : \mathbb{R}^n \rightarrow Y$  by  $A(z) = v_\emptyset + \sum_{i=1}^n z_i v_{\{i\}}$ . For  $z \in [0, 1]^n$  choose  $y \in \frac{1}{2^m}\{0, \dots, 2^m - 1\}^n$  with  $|z_i - y_i| \leq 1/2^{m+1}$  for all  $i \in \{1, \dots, n\}$ . If we assume in addition that  $z \in [0, \sqrt{\varepsilon}]^n$  then

also  $0 \leq y_i \leq \frac{1}{2^{m+1}} + \sqrt{\varepsilon}$  for all  $i \in \{1, \dots, n\}$ . Setting  $g(y) = \sum_{S \subseteq \{1, \dots, n\}} W_S(y) v_S$ , we have

$$\begin{aligned} & \|f(z) - A(z)\|_Y \\ & \leq \|f(z) - f(y)\|_Y + \|f(y) - g(y)\|_Y + \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S| \geq 2}} W_S(y) \|v_S\|_Y + \sum_{i=1}^n |z_i - y_i| \cdot \|v_{\{i\}}\|_Y \\ & \leq \frac{\sqrt{n}}{2^{m+1}} + n\varepsilon + \sum_{k=2}^n \binom{n}{k} \left( \sqrt{\varepsilon} + \frac{1}{2^{m+1}} \right)^k 2^{k-1} + \frac{n}{2^{m+1}} \end{aligned} \quad (30)$$

$$\begin{aligned} & = \frac{\sqrt{n} + n}{2^{m+1}} + n\varepsilon + \frac{1}{2} \left( \left( 1 + 2\sqrt{\varepsilon} + \frac{1}{2^m} \right)^n - 1 - 2n\sqrt{\varepsilon} - \frac{n}{2^m} \right) \\ & \leq 3n\varepsilon + \frac{(1 + \sqrt{5\varepsilon})^n - 1 - n\sqrt{5\varepsilon}}{2} \end{aligned} \quad (31)$$

$$\leq 8n^2\varepsilon, \quad (32)$$

where in (30) we used the fact that  $f$  is 1-Lipschitz and  $\|y - z\|_2 \leq \sqrt{n}\|y - z\|_\infty \leq \frac{\sqrt{n}}{2^{m+1}}$ , the estimates (28), (29), and the above bounds on  $\|y - z\|_\infty$  and  $\|y\|_\infty$ . In (31) we used our assumption  $2^m \geq 2/\varepsilon \geq 10n^2$ , which directly implies that  $2\sqrt{\varepsilon} + 2^{-m} \leq \sqrt{5\varepsilon} \leq 1/n$ , together with the fact that the mapping  $s \mapsto (1 + s)^n - 1 - ns$  is increasing on  $(0, \infty)$ . In (32) we used the elementary inequality  $(1 + s)^n - 1 - ns \leq 2n^2s^2$ , which is valid when  $s \in (0, 1/n)$ .

It remains to prove (28) and (29), which will be done by induction on  $n$ . For  $n = 1$  set  $v_\emptyset = f(0)$  and  $v_{\{1\}} = f(1) - f(0)$ . Since  $f$  is 1-Lipschitz we know that  $\|v_{\{1\}}\|_Y \leq 1$ , proving (28). For the above choices of  $v_\emptyset, v_{\{1\}}$ , the estimate (29) is the same as the assumption  $\mathcal{D}_1^m(f)(x) \leq \varepsilon$  (recall that  $\vartheta = 1$ ).

If  $n > 1$  apply the inductive hypothesis to the functions  $f_0, f_1 : \mathbb{R}^{n-1} \rightarrow Y$  given by  $f_0(y_1, \dots, y_{n-1}) = f(y_1, \dots, y_{n-1}, 0)$  and  $f_1(y_1, \dots, y_{n-1}) = f(y_1, \dots, y_{n-1}, 1)$ . One obtains  $\{v_S^0\}_{S \subseteq \{1, \dots, n-1\}}, \{v_S^1\}_{S \subseteq \{1, \dots, n-1\}} \subseteq Y$  satisfying  $v_\emptyset^0 = f(0)$ ,  $v_\emptyset^1 = f(e_n)$ , for all nonempty  $S \subseteq \{1, \dots, n-1\}$  we have  $\|v_S^0\|_Y, \|v_S^1\|_Y \leq 2^{|S|-1}$ , and if we define  $g_0, g_1 : \mathbb{R}^{n-1} \rightarrow Y$  by

$$g_i(y) \stackrel{\text{def}}{=} \sum_{S \subseteq \{1, \dots, n-1\}} W_S(y) v_S^i,$$

then

$$\max \{\|g_0(y) - f_0(y)\|_Y, \|g_1(y) - f_1(y)\|_Y\} \leq \varepsilon(n-1) \quad (33)$$

for all  $y \in \frac{1}{2^m}\{0, \dots, 2^m\}^{n-1}$ . For  $S \subseteq \{1, \dots, n\}$  define

$$v_S \stackrel{\text{def}}{=} \begin{cases} v_S^0 & \text{if } n \notin S, \\ v_{S \setminus \{n\}}^1 - v_{S \setminus \{n\}}^0 & \text{if } n \in S. \end{cases} \quad (34)$$

So,  $v_\emptyset = v_\emptyset^0 = f(0)$ . If  $S \neq \emptyset$  and  $n \notin S$  then have  $\|v_S\|_Y = \|v_S^0\|_Y \leq 2^{|S|-1}$ . If  $n \in S$  and  $S \setminus \{n\} \neq \emptyset$  then  $\|v_S\|_Y \leq \|v_{S \setminus \{n\}}^0\|_Y + \|v_{S \setminus \{n\}}^1\|_Y \leq 2^{|S|-2} + 2^{|S|-2} = 2^{|S|-1}$ . Finally, since  $f$  is 1-Lipschitz we have  $\|v_{\{n\}}\|_Y = \|v_\emptyset^1 - v_\emptyset^0\|_Y = \|f(e_n) - f(0)\|_Y \leq 1$ . This completes the proof of (28). To prove (29) define for  $y \in \mathbb{R}^n$ ,

$$g(y) \stackrel{\text{def}}{=} \sum_{S \subseteq \{1, \dots, n\}} W_S(y) v_S \stackrel{(34)}{=} (1 - y_n) g_0(y_1, \dots, y_{n-1}) + y_n g_1(y_1, \dots, y_{n-1}). \quad (35)$$

The assumption  $\mathcal{D}_1^m(f)(x) \leq \varepsilon$  implies that for all  $y \in \frac{1}{2^m}\{0, \dots, 2^m\}^{n-1}$  and all  $k \in \{0, \dots, 2^m\}$  we have

$$\left\| f\left(y_1, \dots, y_{n-1}, \frac{k}{2^m}\right) - \left(1 - \frac{k}{2^m}\right) f_0(y) - \frac{k}{2^m} f_1(y) \right\|_Y \leq \varepsilon. \quad (36)$$

Hence,

$$\begin{aligned} & \left\| f\left(y_1, \dots, y_{n-1}, \frac{k}{2^m}\right) - g\left(y_1, \dots, y_{n-1}, \frac{k}{2^m}\right) \right\|_Y \\ & \stackrel{(35) \wedge (36)}{\leq} \varepsilon + \left(1 - \frac{k}{2^m}\right) \|f_0(y) - g_0(y)\|_Y + \frac{k}{2^m} \|f_1(y) - g_1(y)\|_Y \stackrel{(33)}{\leq} \varepsilon n. \end{aligned} \quad (37)$$

Since (37) holds for all  $y \in \frac{1}{2^m}\{0, \dots, 2^m\}^{n-1}$  and all  $k \in \{0, \dots, 2^m\}$ , the proof of (29) is complete.  $\square$

*Proof of Theorem 1.1 when  $n \geq 2$ .* Our goal is to show that if  $(Y, \|\cdot\|_Y)$  is a Banach space satisfying (2) then for every  $\varepsilon \in (0, \frac{1}{2})$  we have

$$r^{X \rightarrow Y}(\varepsilon) \geq R \stackrel{\text{def}}{=} \varepsilon^{K^p n^{20(n+p)}/\varepsilon^{2p+2n-2}}. \quad (38)$$

Assume for contradiction that (38) fails. Then there exists  $\varepsilon \in (0, \frac{1}{2})$  and a  $\frac{1}{\sqrt{n}}$ -Lipschitz function  $f : B_X \rightarrow Y$  such that for all  $\rho \geq R$  and  $y \in X$  such that  $y + \rho B_X \subseteq B_X$ , if  $A : X \rightarrow Y$  is affine then

$$\sup_{z \in y + \rho B_X} \frac{\|f(z) - A(z)\|_Y}{\rho} > \frac{\varepsilon}{\sqrt{n}}. \quad (39)$$

We claim that this implies the following statement.

$$x \in \frac{1}{2}B_X \wedge \vartheta \in \left[\frac{32n^{5/2}}{\varepsilon}R, \frac{1}{2n}\right] \wedge 2^m \in \left[\frac{512n^5}{\varepsilon^2}, \infty\right) \cap \mathbb{N} \implies \mathcal{D}_\vartheta^m(f)(x) > \frac{\varepsilon^2}{256n^5}. \quad (40)$$

Indeed, note that, because  $\|\cdot\| \leq \sqrt{n}\|\cdot\|_2 \leq n\|\cdot\|_\infty$ , the assumptions in (40) imply that  $x + [0, \vartheta]^n \subseteq B_X$ . Since  $f$  is  $\frac{1}{\sqrt{n}}$ -Lipschitz, it is 1-Lipschitz with respect to the Euclidean norm. If  $\mathcal{D}_\vartheta^m(f)(x) \leq \varepsilon^2/(256n^5)$  then it would follow from Lemma 3.1 that there exists an affine mapping  $A : X \rightarrow Y$  such that

$$\sup_{z \in x + [0, \varepsilon\vartheta/(16n^{5/2})]} \frac{\|f(z) - A(z)\|_Y}{\varepsilon\vartheta/(16n^{5/2})} \leq 8n^2 \sqrt{\frac{\varepsilon^2}{256n^5}} = \frac{\varepsilon}{2\sqrt{n}}. \quad (41)$$

Because  $\|\cdot\|_X \geq \|\cdot\|_2$ , we have  $[-1, 1]^n \supseteq B_X$ . Setting  $\rho = \varepsilon\vartheta/(32n^{5/2}) \geq R$ , it follows that  $x + [0, \varepsilon\vartheta/(16n^{5/2})]^n \supseteq y + \rho B_X$  for some  $y \in X$  with  $y + \rho B_X \subseteq B_X$ . Hence (41) contradicts (39), completing the verification of (40).

It remains to argue that (40) leads to a contradiction. To this end, consider the following quantity, defined for every  $x \in \frac{1}{2}B_X$ ,  $m, k \in \mathbb{N} \cup \{0\}$  and  $\vartheta \in (0, 1/(2n)]$ .

$$H_{m,k}^\vartheta(f)(x) \stackrel{\text{def}}{=} \frac{1}{2^{m(n-1)}} \sum_{j=1}^n \sum_{\substack{y \in \{0, \dots, 2^m-1\}^n \\ y_j=0}} E_k^{0,\vartheta} \left( f_j^{x+\vartheta 2^{-m}y} \right). \quad (42)$$

In (42), recall the notation  $f_j^u(t) = f(u + te_j)$  and the definition (13). One checks directly from the definition (42) that the following recursive relation holds true. If  $\alpha, \beta, \gamma \in \mathbb{N} \cup \{0\}$  and  $\alpha \geq \beta$  then for every  $\vartheta \in (0, 1/(2n)]$ ,

$$H_{\alpha, \beta + \gamma}^{\vartheta}(f)(0) = \frac{1}{2^{\beta n}} \sum_{x \in \{0, \dots, 2^{\beta} - 1\}^n} H_{\alpha - \beta, \gamma}^{\vartheta/2^{\beta}}(f) \left( \frac{\vartheta}{2^{\beta}} x \right). \quad (43)$$

Observe that the fact that  $f$  is  $\frac{1}{\sqrt{n}}$ -Lipschitz and  $\|e_j\|_X \leq \sqrt{n}\|e_j\|_2 = \sqrt{n}$  implies that in each of the summands in (42) the function  $f_j^{x+\vartheta 2^{-m}y} : [0, \vartheta] \rightarrow Y$  is 1-Lipschitz. Therefore we have the point-wise bound  $E_k^{0, \vartheta} \left( f_j^{x+\vartheta 2^{-m}y} \right) \leq 1$  for each summand in (42), implying that

$$H_{m, k}^{\vartheta}(f)(x) \leq n. \quad (44)$$

Set

$$m \stackrel{\text{def}}{=} \left\lceil \log_2 \left( \frac{512n^5}{\varepsilon^2} \right) \right\rceil, \quad (45)$$

and

$$M \stackrel{\text{def}}{=} \left\lfloor \frac{1}{m} \log_2 \left( \frac{\varepsilon}{64n^{7/2}R} \right) \right\rfloor. \quad (46)$$

Fix also an integer  $k \in [0, M]$  and set

$$\vartheta \stackrel{\text{def}}{=} \frac{1}{2^{km+1}n}. \quad (47)$$

Then  $\vartheta \geq 32n^{5/2}R/\varepsilon$  (recall (38), (45), (46)). It follows from (40) that  $\mathcal{D}_{\vartheta}^m(f)(x) \geq \varepsilon^2/(2^8n^5)$ . By the definition (26), this means that there exists  $j \in \{1, \dots, n\}$  and  $w \in x + \vartheta F_j^m$  (recall that  $F_j^m = \{z \in \frac{1}{2^m}\{0, \dots, 2^m\}^n : z_j = 0\}$ ), such that for some  $s \in \{0, \dots, 2^m\}$  we have

$$\left\| f \left( w + \frac{s\vartheta}{2^m} e_j \right) - f(w) - \frac{s}{2^m} (f(w + \vartheta e_j) - f(w)) \right\|_Y \geq \frac{\varepsilon^2}{2^9 n^6 2^{km}}. \quad (48)$$

Denote  $\ell = (M + 1 - k)m$  and consider the set

$$C \stackrel{\text{def}}{=} \left\{ y \in \{0, \dots, 2^{\ell} - 1\} : y_j = 0 \wedge \left\| y - \frac{2^{\ell}}{\vartheta} (w - x) \right\|_{\infty} \leq \frac{\varepsilon^2 2^{\ell}}{2^{10} n^{11/2}} \right\}. \quad (49)$$

Then

$$|C| \geq \left\lfloor \frac{\varepsilon^2 2^{\ell}}{2^{10} n^{11/2}} \right\rfloor^{n-1} \geq \left( \frac{\varepsilon^2}{2^{11} n^{11/2}} \right)^{n-1} 2^{\ell(n-1)}. \quad (50)$$

Since the Lipschitz constant of  $f$  with respect to the  $\ell_{\infty}$  norm is at most  $\sqrt{n}$ , it follows from (48) that for every  $y \in C$  we have

$$\begin{aligned} & \left\| f \left( x + \frac{\vartheta}{2^{\ell}} y + \frac{s\vartheta}{2^m} e_j \right) - f \left( x + \frac{\vartheta}{2^{\ell}} y \right) - \frac{s}{2^m} \left( f \left( x + \frac{\vartheta}{2^{\ell}} y + \vartheta e_j \right) - f \left( x + \frac{\vartheta}{2^{\ell}} y \right) \right) \right\|_Y \\ & \stackrel{(49)}{\geq} \frac{\varepsilon^2}{2^9 n^6 2^{km}} - 2\sqrt{n} \cdot \frac{\vartheta}{2^{\ell}} \cdot \frac{\varepsilon^2 2^{\ell}}{2^{10} n^{11/2}} \stackrel{(47)}{=} \frac{\varepsilon^2}{2^{10} n^6 2^{km}}. \end{aligned} \quad (51)$$

An equivalent way to write (51) is as follows.

$$\left\| f_j^{x+\vartheta 2^{-\ell}y} \left( \frac{s}{2^m} \right) - L_{f_j^{x+\vartheta 2^{-\ell}y}}^{0, \vartheta} \left( \frac{s}{2^m} \right) \right\|_Y \geq \frac{\varepsilon^2}{2^{10} n^6 2^{km}}.$$

An application of Lemma 2.1 now implies that for every  $y \in C$  we have

$$\forall y \in C, \quad E_m^{0,\vartheta} \left( f_j^{x+\vartheta 2^{-\ell}y} \right) \geq \frac{\left\| f_j^{x+\vartheta 2^{-\ell}y}(\vartheta) - f_j^{x+\vartheta 2^{-\ell}y}(0) \right\|_Y^p}{\vartheta^p} + \left( \frac{\varepsilon^2}{K(4n)^5} \right)^p, \quad (52)$$

where  $K$  is the constant in (2). Also, by convexity (see (16)), for every  $i \in \{1, \dots, n\}$  we have

$$y \in \{0, \dots, 2^\ell\} \wedge y_i = 0 \implies E_m^{0,\vartheta} \left( f_i^{x+\vartheta 2^{-\ell}y} \right) \geq \frac{\left\| f_i^{x+\vartheta 2^{-\ell}y}(\vartheta) - f_i^{x+\vartheta 2^{-\ell}y}(0) \right\|_Y^p}{\vartheta^p}. \quad (53)$$

Hence,

$$\begin{aligned} H_{\ell,m}^\vartheta(f)(x) &\stackrel{(42)}{=} \frac{1}{2^{\ell(n-1)}} \sum_{y \in C} E_m^{0,\vartheta} \left( f_j^{x+\vartheta 2^{-\ell}y} \right) + \frac{1}{2^{\ell(n-1)}} \sum_{\substack{y \in \{0, \dots, 2^\ell-1\}^n \\ y_j=0 \\ y \notin C}} E_m^{0,\vartheta} \left( f_j^{x+\vartheta 2^{-\ell}y} \right) \\ &\quad + \frac{1}{2^{\ell(n-1)}} \sum_{\substack{i \in \{1, \dots, n\} \\ i \neq j}} \sum_{\substack{y \in \{0, \dots, 2^\ell-1\}^n \\ y_i=0}} E_m^{0,\vartheta} \left( f_i^{x+\vartheta 2^{-\ell}y} \right) \\ &\stackrel{(52) \wedge (53)}{\geq} \frac{1}{2^{\ell(n-1)}} \sum_{i=1}^n \sum_{\substack{y \in \{0, \dots, 2^\ell-1\}^n \\ y_i=0}} \frac{\left\| f_i^{x+\vartheta 2^{-\ell}y}(\vartheta) - f_i^{x+\vartheta 2^{-\ell}y}(0) \right\|_Y^p}{\vartheta^p} + \frac{|C|}{2^{\ell(n-1)}} \left( \frac{\varepsilon^2}{K(4n)^5} \right)^p \\ &\stackrel{(50)}{\geq} \frac{1}{2^{\ell(n-1)}} \sum_{i=1}^n \sum_{\substack{y \in \{0, \dots, 2^\ell-1\}^n \\ y_i=0}} \frac{\left\| f \left( x + \frac{\vartheta}{2^\ell} (y + 2^\ell e_i) \right) - f \left( x + \frac{\vartheta}{2^\ell} y \right) \right\|_Y^p}{\vartheta^p} + \frac{\varepsilon^{2(n-1+p)}}{K^p (4n)^{6n+5p}}. \end{aligned} \quad (54)$$

Now, using the recursive identity (43), we have

$$H_{(M+1)m, (k+1)m}^{1/(2n)}(f)(0) = \frac{1}{2^{kmn}} \sum_{x \in \{0, \dots, 2^{km}-1\}^n} H_{(M+1-k)m, m}^{2^{-km}/(2n)}(f) \left( \frac{2^{-km}}{2n} x \right). \quad (55)$$

We relate (55) to (54) by noting the following identity, in which we recall that  $\vartheta$  is given in (47) and  $\ell = (M+1-k)m$ .

$$\begin{aligned} &\frac{1}{2^{kmn+\ell(n-1)}} \sum_{x \in \{0, \dots, 2^{km}-1\}^n} \sum_{i=1}^n \sum_{\substack{y \in \{0, \dots, 2^\ell-1\}^n \\ y_i=0}} \frac{\left\| f \left( \frac{2^{-km}}{2n} x + \frac{\vartheta}{2^\ell} (y + 2^\ell e_i) \right) - f \left( \frac{2^{-km}}{2n} x + \frac{\vartheta}{2^\ell} y \right) \right\|_Y^p}{\vartheta^p} \\ &= \frac{1}{2^{(M+1)m(n-1)+km}} \sum_{i=1}^n \sum_{z \in \{0, \dots, 2^{(M+1)m}-1\}^n} \frac{\left\| f \left( \frac{2^{-(M+1)m}}{2n} z + \frac{2^{-km}}{2n} e_i \right) - f \left( \frac{2^{-(M+1)m}}{2n} z \right) \right\|_Y^p}{(2^{-km}/(2n))^p} \\ &\stackrel{(13)}{=} \frac{1}{2^{(M+1)m(n-1)}} \sum_{i=1}^n \sum_{\substack{y \in \{1, \dots, 2^{(M+1)m}-1\} \\ y_i=0}} E_{km}^{0,1/(2n)} \left( f_i^{\frac{2^{-(M+1)m}}{2n} y} \right) \\ &\stackrel{(42)}{=} H_{(M+1)m, km}^{1/(2n)}(f)(0). \end{aligned} \quad (56)$$

By combining (54), (55) and (56) we conclude that

$$\forall k \in \{0, \dots, M\}, \quad H_{(M+1)m, (k+1)m}^{1/(2n)}(f)(0) \geq H_{(M+1)m, km}^{1/(2n)}(f)(0) + \frac{\varepsilon^{2(n-1+p)}}{K^p(4n)^{6n+5p}}.$$

Hence,

$$n \stackrel{(44)}{\geq} H_{(M+1)m, (M+1)m}^{1/(2n)}(f)(0) \geq (M+1) \frac{\varepsilon^{2(n-1+p)}}{K^p(4n)^{6n+5p}}. \quad (57)$$

Recalling the definitions (38), (45) and (46), and that  $K \geq 1$ ,  $\varepsilon \in (0, \frac{1}{2})$  and  $p, n \geq 2$ , one checks that (57) is a contradiction.  $\square$

#### 4. AN EXAMPLE

We start with a simple one dimensional construction.

**Lemma 4.1.** *Fix  $p \in [2, \infty)$  and  $m \in \mathbb{N}$ . There exists a 1-Lipschitz function  $f : [0, 1] \rightarrow \ell_p^m$  with  $f(0) = f(1) = 0$  such that for every  $0 \leq a < b \leq 1$  with  $b - a \geq 4/2^m$  and every affine mapping  $A : \mathbb{R} \rightarrow \ell_p^m$  we have*

$$\sup_{x \in [a, b]} \frac{\|f_m(x) - A(x)\|_p}{(b-a)/2} > \frac{1}{8m^{1/p}}.$$

Consequently, if we set  $\varepsilon = \frac{1}{8m^{1/p}}$  then

$$r^{\mathbb{R} \rightarrow \ell_p}(\varepsilon) \leq \frac{4}{2^{1/(8\varepsilon)^p}}.$$

*Proof.* Define inductively a sequence of functions  $\{f_k : [0, 1] \rightarrow \ell_p^m\}_{k=0}^m$  as follows. Let  $\{e_1, \dots, e_m\}$  be the standard basis of  $\ell_p^m$ . Set  $f_0 \equiv 0$ . Assume that  $k \in \mathbb{N}$  and we have defined  $f_{k-1}$  to be affine on each of the dyadic intervals  $\{[j/2^{k-1}, (j+1)/2^{k-1}]\}_{j=0}^{2^{k-1}-1}$ . For every  $j \in \{0, \dots, 2^{k-1}\}$  define  $f_k(j/2^{k-1}) = f_{k-1}(j/2^{k-1})$  and

$$f_k\left(\frac{2j+1}{2^k}\right) = f_{k-1}\left(\frac{2j+1}{2^k}\right) + \frac{1}{m^{1/p}2^k}e_k. \quad (58)$$

Let  $f_k$  be the piecewise affine extension of the above values of  $f_k$  on  $\{j/2^k\}_{j=0}^{2^k-1}$ . A straightforward induction shows that

$$\left\| f_k\left(\frac{j+1}{2^k}\right) - f_k\left(\frac{j}{2^k}\right) \right\|_p = \frac{1}{2^k} \left(\frac{k}{m}\right)^{1/p}.$$

Thus  $f_m$  is 1-Lipschitz.

Assume for contradiction that  $0 \leq a < b \leq 1$  satisfy  $b - a \geq 4/2^m$ , and there exists an affine mapping  $A : \mathbb{R} \rightarrow \ell_p^m$  such that

$$\sup_{x \in [a, b]} \frac{\|f_m(x) - A(x)\|_p}{(b-a)/2} \leq \frac{1}{8m^{1/p}}. \quad (59)$$

There exists  $k \in \{1, \dots, m\}$  such that  $4/2^k \leq b - a < 8/2^k$ . Because  $b - a \geq 4/2^k$  there is  $j \in \{0, \dots, 2^{k-1} - 1\}$  such that  $[j/2^{k-1}, (j+1)/2^{k-1}] \subseteq [a, b]$ . Now, since  $A$  is affine and  $f_{k-1}$

is affine on  $[j/2^{k-1}, (j+1)/2^{k-1}]$ ,

$$\begin{aligned} \frac{b-a}{8m^{1/p}} &\stackrel{(59)}{\geq} \left\| f_m \left( \frac{j/2^{k-1} + (j+1)/2^{k-1}}{2} \right) - \frac{f_m(j/2^{k-1}) + f_m((j+1)/2^{k-1})}{2} \right\|_p \\ &= \left\| f_k \left( \frac{2j+1}{2^k} \right) - f_{k-1} \left( \frac{2j+1}{2^k} \right) \right\|_p \stackrel{(58)}{=} \frac{1}{m^{1/p} 2^k}, \end{aligned}$$

in contradiction to the fact that  $b-a < 8/2^k$ .  $\square$

**Lemma 4.2.** Fix  $p \in [2, \infty)$  and  $m, n \in \mathbb{N}$ . There exists a 1-Lipschitz function  $g : \mathbb{R} \rightarrow \ell_p^{m+1}$  such that for every  $y \in \mathbb{R}$  with  $|y| \leq 1/\sqrt{n}$ , every  $r \geq 32/(\sqrt{n}2^m)$ , and every affine mapping  $A : \mathbb{R} \rightarrow \ell_p^{m+1}$ ,

$$\sup_{x \in [y-r, y+r]} \frac{\|g(x) - A(x)\|_p}{r} > \frac{1}{16m^{1/p}}.$$

*Proof.* Let  $\{e_1, \dots, e_{m+1}\}$  denote the standard basis of  $\ell_p^{m+1}$ . Define  $g : \mathbb{R} \rightarrow \ell_p^{m+1}$  by

$$g(x) \stackrel{\text{def}}{=} \begin{cases} \frac{4}{\sqrt{n}} f \left( \frac{\sqrt{n}}{4} x + \frac{1}{2} \right) & \text{if } |x| \leq \frac{2}{\sqrt{n}}, \\ \left( |x| - \frac{2}{\sqrt{n}} \right) e_{m+1} & \text{otherwise.} \end{cases}$$

where  $f = f_m : [0, 1] \rightarrow \ell_p^m = \text{span}(\{e_1, \dots, e_m\})$  is the function from Lemma 4.1. Because  $f$  is 1-Lipschitz and  $f(0) = f(1) = 0$ , one checks that  $g$  is 1-Lipschitz.

Fix an affine mapping  $A : \mathbb{R} \rightarrow \ell_p$  and take  $y \in \mathbb{R}$  satisfying  $|y| \leq 1/\sqrt{n}$ . Suppose that  $r \geq 32/(\sqrt{n}2^m)$ . If in addition  $r \leq 8/\sqrt{n}$  then write  $[y-r, y+r] \cap [-2/\sqrt{n}, 2/\sqrt{n}] = [a, b]$ , where  $b-a \geq r/2 \geq 16/(\sqrt{n}2^m)$ . By Lemma 4.1,

$$\begin{aligned} \sup_{x \in [y-r, y+r]} \frac{\|g(x) - A(x)\|_p}{r} &\geq \sup_{x \in [a, b]} \frac{\left\| \frac{4}{\sqrt{n}} f \left( \frac{\sqrt{n}}{4} x + \frac{1}{2} \right) - A(x) \right\|_p}{2(b-a)} \\ &= \frac{1}{2} \sup_{z \in \left[ \frac{\sqrt{n}}{4} a + \frac{1}{2}, \frac{\sqrt{n}}{4} b + \frac{1}{2} \right]} \frac{\left\| f(z) - \frac{\sqrt{n}}{4} A \left( \frac{4}{\sqrt{n}} z - \frac{2}{\sqrt{n}} \right) \right\|_p}{\left( \frac{\sqrt{n}}{4} b - \frac{\sqrt{n}}{4} a \right) / 2} > \frac{1}{16m^{1/p}}. \end{aligned}$$

It remains to deal with the case  $r > 8/\sqrt{n}$ . In this case  $y-r, y+r \notin [-2/\sqrt{n}, \sqrt{n}]$ , so

$$\langle g(y \pm r), e_{m+1} \rangle = |y \pm r| - \frac{2}{\sqrt{n}} \geq r - |y| - \frac{2}{\sqrt{n}} \geq r - \frac{3}{\sqrt{n}} > \frac{5}{\sqrt{n}} > 0.$$

Assume for contradiction that  $\|g(x) - A(x)\|_p \leq r/(16m^{1/p})$  for all  $x \in [y-r, y+r]$ . Then, since  $A$  is affine,

$$\langle A(y), e_{m+1} \rangle = \frac{\langle A(y+r), e_{m+1} \rangle + \langle A(y-r), e_{m+1} \rangle}{2} \geq r - \frac{3}{\sqrt{n}} - \frac{r}{16m^{1/p}}.$$

Hence,

$$\langle g(y), e_{m+1} \rangle \geq \langle A(y), e_{m+1} \rangle - \|g(y) - A(y)\|_p \geq \left( 1 - \frac{1}{8m^{1/p}} \right) r - \frac{3}{\sqrt{n}} > \frac{r}{2} - \frac{3}{\sqrt{n}} > \frac{1}{\sqrt{n}},$$

contradicting the fact that, since  $|y| \leq 2/\sqrt{n}$ , we have  $\langle g(y), e_{m+1} \rangle = 0$ .  $\square$

We now use the function  $g$  of Lemma 4.2 as a building block of a function  $F : \ell_2^n \rightarrow \ell_2^n(\ell_p^{m+1})$  whose affine approximability properties deteriorate with the dimension  $n$ . This step is similar to an argument in the proof of Theorem 2.7 in [2].

**Lemma 4.3.** *For every  $p \in [2, \infty)$  and every  $m, n \in \mathbb{N}$  there exists a 1-Lipschitz function  $F : \ell_2^n \rightarrow \ell_2^n(\ell_p^{m+1})$  such that for every*

$$r \geq \frac{32}{\sqrt{n}2^m}$$

*and every affine mapping  $A : \ell_2^n \rightarrow \ell_2^n(\ell_p^{m+1})$ ,*

$$\sup_{x \in y + rB_{\ell_2^n}} \frac{\|F(x) - A(x)\|_{\ell_2^n(\ell_p^{m+1})}}{r} > \frac{1}{16m^{1/p}}.$$

*Consequently, if we set  $\varepsilon = \frac{1}{16m^{1/p}}$  then for  $X = \ell_2^n$  and  $Y = \ell_2^n(\ell_p^{m+1})$ ,*

$$r^{X \rightarrow Y}(\varepsilon) \leq \frac{32}{\sqrt{n}2^{1/(16\varepsilon)^p}}.$$

*Proof.* Let  $g : \mathbb{R} \rightarrow \ell_p^{m+1}$  be the function from Lemma 4.2. Since  $g$  is 1-Lipschitz, if we define  $F(x_1, \dots, x_n) = (g(x_1), \dots, g(x_n))$  then  $F : \ell_2^n \rightarrow \ell_2^n(\ell_p^{m+1})$  is 1-Lipschitz. Fixing  $r \geq 32/(\sqrt{n}2^m)$ , suppose that  $A : \ell_2^n \rightarrow \ell_2^n(\ell_p^{m+1})$  is affine and  $y + rB_{\ell_2^n} \subseteq B_{\ell_2^n}$ . Since  $\|y\|_2 < 1$ , there exists  $i \in \{1, \dots, n\}$  such that  $|y_i| < 1/\sqrt{n}$ . Writing  $A(x) = (A_1(x), \dots, A_n(x))$ , define  $A'_i : \mathbb{R} \rightarrow \ell_p^{m+1}$  by  $A'_i(t) = A_i\left(\sum_{j \in \{1, \dots, n\} \setminus \{i\}} y_j e_j + t e_i\right)$ . By Lemma 4.2 we know that

$$\sup_{x \in y + rB_{\ell_2^n}} \frac{\|F(x) - A(x)\|_{\ell_2^n(\ell_p^{m+1})}}{r} \geq \sup_{t \in [y_i - r, y_i + r]} \frac{\|g(t) - A'_i(t)\|_p}{r} > \frac{1}{16m^{1/p}},$$

where we used that fact that  $y + [-r, r]e_i \subseteq y + rB_{\ell_2^n}$ . □

## REFERENCES

- [1] K. Ball, E. A. Carlen, and E. H. Lieb. Sharp uniform convexity and smoothness inequalities for trace norms. *Invent. Math.*, 115(3):463–482, 1994.
- [2] S. Bates, W. Johnson, J. Lindenstrauss, D. Preiss, and G. Schechtman. Affine approximation of Lipschitz functions and nonlinear quotients. *Geometric and Functional Analysis*, 9:1092–1127, 1999.
- [3] Y. Benyamini. The uniform classification of Banach spaces. In *Texas functional analysis seminar 1984–1985 (Austin, Tex.)*, Longhorn Notes, pages 15–38. Univ. Texas Press, Austin, TX, 1985.
- [4] Y. Benyamini and J. Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000.
- [5] J. Bourgain. Remarks on the extension of Lipschitz maps defined on discrete sets and uniform homeomorphisms. In *Geometrical aspects of functional analysis (1985/86)*, volume 1267 of *Lecture Notes in Math.*, pages 157–167. Springer, Berlin, 1987.
- [6] D. L. Burkholder. Martingales and singular integrals in Banach spaces. In *Handbook of the geometry of Banach spaces, Vol. I*, pages 233–269. North-Holland, Amsterdam, 2001.
- [7] J. Cheeger, B. Kleiner, and A. Naor. Compression bounds for Lipschitz maps from the Heisenberg group to  $L_1$ . Preprint available at <http://arxiv.org/abs/0910.2026>. To appear in *Acta Math.*, 2009.
- [8] J. Diestel, H. Jarchow, and A. Tonge. *Absolutely summing operators*, volume 43 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.



- [9] P. Enflo. Banach spaces which can be given an equivalent uniformly convex norm. In *Proceedings of the International Symposium on Partial Differential Equations and the Geometry of Normed Linear Spaces (Jerusalem, 1972)*, volume 13, pages 281–288 (1973), 1972.
- [10] A. Eskin, D. Fisher, and K. Whyte. Quasi-isometries and rigidity of solvable groups. *Pure Appl. Math. Q.*, 3(4, part 1):927–947, 2007.
- [11] T. Figiel. On the moduli of convexity and smoothness. *Studia Math.*, 56(2):121–155, 1976.
- [12] I. Gelfand. Abstrakte Funktionen und lineare Operatoren. *Rec. Math. Moscou, n. Ser.*, 4:235–284, 1938.
- [13] O. Giladi, A. Naor, and G. Schechtman. Bourgain’s discretization theorem. Preprint available at <http://arxiv.org/abs/1110.5368>, 2011.
- [14] S. Heinrich and P. Mankiewicz. Applications of ultrapowers to the uniform and Lipschitz classification of Banach spaces. *Studia Math.*, 73(3):225–251, 1982.
- [15] T. Hytönen, S. Li, and A. Naor. Quantitative affine approximation for UMD targets. Preprint, 2011.
- [16] R. C. James. Uniformly non-square Banach spaces. *Ann. of Math. (2)*, 80:542–550, 1964.
- [17] R. C. James. Super-reflexive Banach spaces. *Canad. J. Math.*, 24:896–904, 1972.
- [18] F. John. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday*, pages 187–204. Interscience Publishers Inc., 1948.
- [19] W. B. Johnson, J. Lindenstrauss, and G. Schechtman. Extensions of Lipschitz maps into Banach spaces. *Israel J. Math.*, 54(2):129–138, 1986.
- [20] A. Naor. Geometric problems in non-linear functional analysis. Master’s thesis, Hebrew University, 1998.
- [21] G. Pisier. Martingales with values in uniformly convex spaces. *Israel J. Math.*, 20(3-4):326–350, 1975.

COURANT INSTITUTE, NEW YORK UNIVERSITY, NEW YORK NY 10012  
*E-mail address:* `seanli@cims.nyu.edu`

COURANT INSTITUTE, NEW YORK UNIVERSITY, NEW YORK NY 10012  
*E-mail address:* `naor@cims.nyu.edu`